

AD _____

GRANT NUMBER DAMD17-96-1-6128

TITLE: Mammogram Screening by Automated Followup: A Feasibility Study

PRINCIPAL INVESTIGATOR: Dragana Brzakovic, Ph.D.

CONTRACTING ORGANIZATION: Lehigh University
Bethlehem, Pennsylvania 18015-3046

REPORT DATE: July 1997

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19970910 073

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

| | | | | |
|---|---|--|---|--|
| 1. AGENCY USE ONLY (Leave blank) | | 2. REPORT DATE July 1997 | 3. REPORT TYPE AND DATES COVERED Annual (15 Jun 96 - 14 Jun 97) | |
| 4. TITLE AND SUBTITLE Mammogram Screening by Automated Followup: A Feasibility Study | | | 5. FUNDING NUMBERS DAMD17-96-1-6128 | |
| 6. AUTHOR(S) Dragana Brzakovic, Ph.D. | | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lehigh University Bethlehem, PA 18015-3046 | | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command Fort Detrick, Frederick, Maryland 21702-5012 | | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER | |
| 11. SUPPLEMENTARY NOTES | | | | |
| 12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited | | | 12b. DISTRIBUTION CODE | |
| 13. ABSTRACT (Maximum 200) This paper describes part of a study aimed at developing a computer-based aid for mammogram screening that makes a detailed comparison between mammograms of the same patient acquired at different screenings and detects changes indicative of cancer. The focus is on determining control points in two mammograms; these points are used to put two mammograms into correspondence. The proposed procedure starts by extracting candidate points in each mammogram at the locations where elongated anatomical structures, such as ducts and blood vessels, intersect. This is achieved by using two multi-scale directional operators and logical operations applied to their combined outputs. The correspondence between points in a mammogram pair is established using an accumulator matrix and similarity criterion applied to each candidate pair. Two similarity criteria are considered. The first measures similarity between the geometry of intersecting elongated structures. The second measures similarity between texture patterns around candidate points. The methods were evaluated on synthetic images and archived film. Evaluation on the synthetic images is summarized in the form of ROC curves. Performance on real mammograms was evaluated by visual inspection involving two untrained observers and an experienced radiologist. A high degree of agreement was found between the observers and the algorithm, and in particular between the radiologist and the algorithm (the agreement was 91%). | | | | |
| 14. SUBJECT TERMS Mammography Digital mammography, mammogram comparison, follow-up, control points, mammogram registration, regional comparison | | | 15. NUMBER OF PAGES 35 | |
| | | | 16. PRICE CODE | |
| | | | DTIC QUALITY INSPECTED 4 | |
| 17. SECURITY CLASSIFICATION OF REPORT Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified | 20. LIMITATION OF ABSTRACT Unlimited | |

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material. /

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.


PI  Signature 7/9/97
Date

Table of Contents

| | |
|--|-----------|
| Front Cover | 1 |
| Standard Form | 2 |
| Foreword | 3 |
| Table of Contents | 4 |
| 1 Introduction | 6 |
| 2 Approach | 7 |
| 3 Extraction of potential control points | 9 |
| 3.1 Rationale for the proposed algorithm | 9 |
| 3.2 Detection of the intersection points | 10 |
| 3.3 Parameter selection | 11 |
| 4 Establishing control point correspondence | 11 |
| 4.1 Correspondence algorithm | 13 |
| 4.2 Identifying reference points | 15 |
| 4.3 Similarity measures and criteria | 15 |
| 4.3.1 Signatures | 15 |
| 4.3.2 Texture characterization | 16 |
| 5 Evaluation: Synthetic images | 17 |
| 5.1 Material selection | 17 |
| 5.2 Evaluation: detection algorithm | 18 |
| 5.2.1 Experiment description | 18 |
| 5.2.2 Results | 18 |
| 5.3 Evaluation: correspondence algorithm | 21 |
| 5.3.1 Experiment description | 22 |
| 5.3.2 Results | 22 |
| 6 Evaluation: Archived film | 23 |
| 6.1 Material selection | 23 |
| 6.2 Experiment description | 23 |
| 6.2.1 Analysis of consistency of human performance | 25 |

| | | |
|----------|---|-----------|
| 6.3 | Results: signatures | 28 |
| 6.4 | Results: Laws' texture measures | 31 |
| 7 | Conclusions | 32 |

1 Introduction

The long-range goal of this work is to couple computer technology and human expertise with the objective to improve accuracy and consistency of mammogram readings while reducing the cost. Specifically, this work is concerned with the problem of detecting early cancerous changes by comparing temporal sequences of mammograms of a same patient. An advantage of considering mammogram sequences relative to analyzing a single mammogram is that it involves comparison, using the older mammogram as a reference, and thus, is likely to reduce the high rate of false positives associated with many computer-based methods. More importantly, following subtle changes may provide for very early cancer detection. However, automating mammogram sequence analysis is a very complex task, and it requires solving a number of subproblems, including standardizing screening procedures, registering mammograms, and characterizing minor changes in texture patterns. The present work focuses on the initial step of mammogram registration. The specific goals of the on-going study are to develop and test digital image processing methods that register mammograms (acquired in regular screenings) and provide for regional mammogram comparison with the objective of detecting early cancerous changes. These methods are envisioned as an integral part of a computer-based aid for mammogram screening which draws the attention of medical experts to suspicious regions in mammograms.

The problem of mammogram registration was encountered early in digital mammography, initially in attempts to compare the same view of the two breasts and search for asymmetries [31]. Some researchers have registered mammogram pairs under the assumption that they are related by translation [6], [23]. Alternatively, registration was performed either manually [29], using control points extracted from the breast outlines [11], or by using optimization procedures involving rotation [25], combined rotation and translation [30], or 2-D unwarping [20]. The view taken in this work is that the precise mammogram registration is intractable. This is due to the fact that these images correspond to compressed, elastic 3-D objects and the images differ primarily due to the fact that there are variations in positioning and compression. These variations are in essence changes in viewpoint, and 2-D transformations can not counteract 3-D changes in viewpoint. This work concentrates on developing an alternative to precise registration by defining corresponding regions in two images, as is done by medical experts. These regions can then be analyzed for changes indicative of cancer. The first step in defining regions is establishing landmarks in a mammogram pair and establishing their correspondence. An important advantage of this approach is that it avoids interpolation inherent in detailed registration and thus provides for meaningful comparison between temporally spaced screenings.

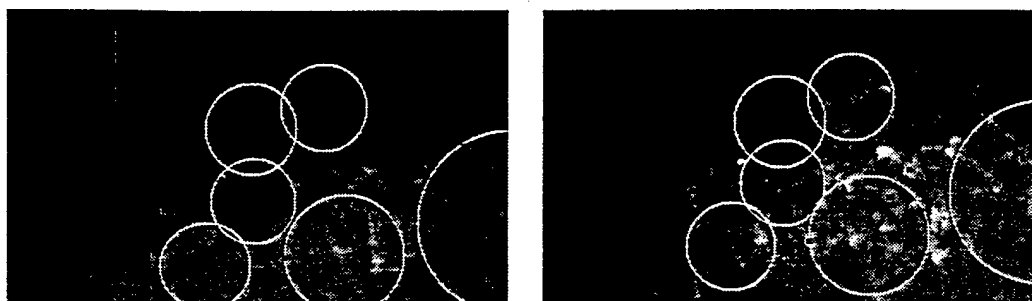
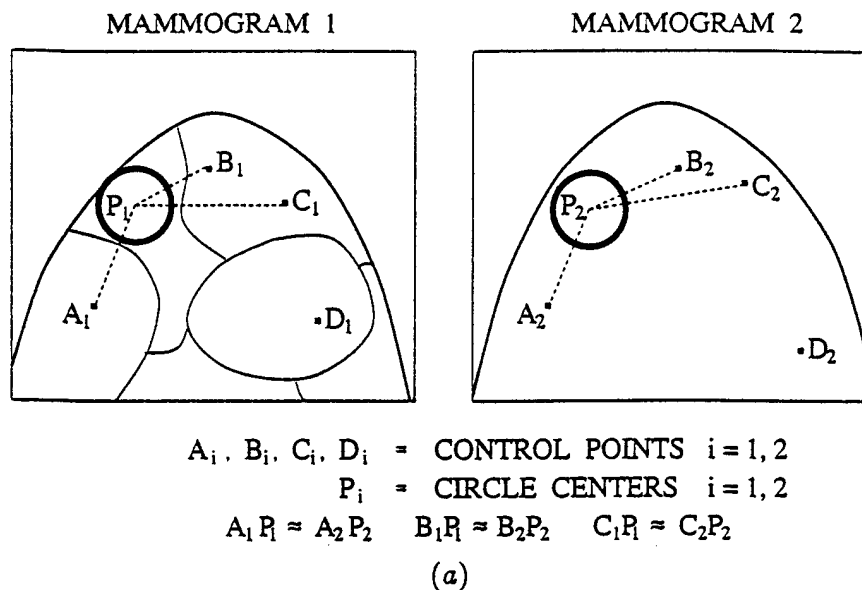
The focus of the work in the past year has been on the problem of extracting the landmark points, also referred to as the potential control points, and establishing their

correspondence. Some aspects of this work are described in our recent publications [27], [28]. Detecting and matching landmark points is resolved in two steps. First, each image is analyzed independently and two sets of landmark points are identified. An intersection of elongated structures is considered to be a landmark point. Next, correspondence between a common subset of landmarks is established by using accumulator matrices which tally all possible matches. The likely matches are determined using similarity criteria. Two criteria are considered: one measuring similarity in the geometry of the elongated structures and the other measuring local textural characteristics. The algorithms were evaluated first on synthetic images where ROC curves were used to study performance under controlled conditions. Next, the algorithms were applied to 31 pairs of mammograms, and the results were visually evaluated by three observers.

The overall approach is described in Section 2. The algorithm for extraction of potential control points is detailed in Section 3, and the matching procedure using both criteria is detailed in Section 4. The evaluation protocols and results are described in Sections 5 and 6, corresponding to synthetic images and archived film, respectively.

2 Approach

The premise behind the solution to the registration problem proposed in this work is that there is no rational basis to develop registration techniques that counter-affect 3-D positioning differences by 2-D image registration. In view of this, we propose an alternative approach which surmounts the problem of precise mammogram registration and performs regional registration instead. In this approach, mammograms are compared regionally, and each pair of regions is defined based on location relative to a set of control points. In order to reduce the effects of variations in breast tissue density on the subsequent analysis, the regions are selected such that each region contains primarily a single tissue type. For this purpose, we utilize image partitioning and constrain regions to belong to a single partition. In order to compensate for 2-D rotational distortions arising from digitization, we match circular regions. Each circle is defined by the location of its center relative to the control points and its radius, which is determined by segmentation of the older mammogram. Both mammograms are covered, without gaps, by corresponding overlapping circles. The procedure used to define circular regions is illustrate in Figure 1(a), and an example of obtained regions is shown in Figure 1(b). The segmentation of the older mammogram, shown in Figure 1(c), is based on our earlier work described in [2]. Regions are compared in terms of their intensity statistics and texture characteristics. The algorithm for mammogram covering is presently under development.



(b)



(c)

Figure 1: Regional correspondence between mammograms: (a) illustration of procedure, in this case the centers of the regions are determined by distances relative to the three closest control points, (b) examples of regions determined following this procedure and using control points established using method described in this work and segmentation of the older mammogram, (c) segmentation of the older mammogram used to determine region extent.

3 Extraction of potential control points

Considering that mammograms are weakly-structured texture patterns and that the relationship between two images cannot be modeled, establishing regional correspondence requires at least determining a set of landmarks common to both images. These landmarks can be used as reference points in establishing locations and geometry of corresponding regions. An approach to defining regions is described in [27], whereby corresponding circular regions are formed by defining centers based on distance from the control points and radii based on tissue characteristics.

Based on the fact that the only distinct elements appearing in practically any mammogram are elongated structures, the most prominent of which correspond to milk ducts and blood vessels, we select the intersections between these structures to be the landmarks. After extracting their intersections in both images, we establish correspondence between the common subset of points. In order to make establishing correspondence practical, we consider only the intersections that are the most likely to appear in both mammograms; consequently, we limit the extraction to the intersections between the most prominent elongated structures. The obtained points are referred to as the potential control points. Locations of the potential control points may vary in two images due to minor positioning differences and changes in compression; however, these variations are in most cases minor and this study concentrates on such cases.

3.1 Rationale for the proposed algorithm

The study of the intensity profiles of elongated structures, [4], shows that the intensity, $i(x)$, changes across the structure cross-section as

$$i(x) = C_1 - C_2\sqrt{a^2 - x^2}, \quad -a \leq x \leq a, \quad (1)$$

where a is the width of the cross-section; constant C_1 models the effects of film, anti-scatter grid and x-ray tube; and constant C_2 models compression and attenuation effects. Effectively, Equation (1) shows that the cross-sections of elongated structures are characterized by roof edge profiles. Consequently, extraction of the elongated structures is, in essence, a task of identifying peaks of the roof edge profiles. Similar tasks are encountered in other applications, e.g., in fingerprint analysis, line drawing understanding, and digital angiography. Many researchers have investigated this problem and have proposed appropriate methods, e.g., [5], [9], [19]. The emphasis of these approaches is on detailed image analysis, extraction of every roof edge with the objective of retaining continuity of the structures. Mammograms are complex images with many intensity variations; thus, extraction of every roof edge yields overwhelming detail. Instead, we propose an alternative algorithm which identifies

only the most prominent roof edges. Specifically, we limit the identification to the widest roof edge profiles.

3.2 Detection of the intersection points

The intersections are detected by first identifying the elongated structures using the modified monotony operator. Prior to applying the operator, images are smoothed using the multi-pass Gaussian filter. The filter is approximated by a 5×5 kernel, as described in [3]. Convolution of a Gaussian function $g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2})$ with Equation (1) gives the new intensity profile i_n of form

$$i_n(x) = i(x) * g(x) = \int_{-1}^{+1} (1 - \sqrt{1 - t^2}) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-t)^2}{2\sigma^2}} dt. \quad (2)$$

(Equation (2) is shown in a simplified form since the objective here is only to discuss the form of smoothed roof edge profiles.) Obtaining values for i_n requires numerical computation, but the form of the equation indicates that the shape of the profile is retained. Therefore, smoothing removes minor intensity variations and does not in effect change the intensity profiles. The effects of smoothing are discussed in detail in Section 5.

Prominent roof edges are extracted using the modified monotony operator. Originally, the monotony operator was proposed in [10] to identify points which were used to derive displacement vector fields in image sequences. The monotony operator compares the gray level of an image pixel with that of its eight neighbors and assigns a value of 0-8 to it, according to the number of neighbors that have a smaller gray value than that of the central pixel. Points of interest, such as edges and corners, are determined by selecting pixels which are assigned specific values. In this work, the concept of the monotony operator is modified as follows:

- The modified monotony operator considers two concentric neighborhoods of different size, a large neighborhood $m \times n$ and a small neighborhood $k \times l$.
- The modified monotony operator assigns a value to an image pixel by counting the pixels in the large neighborhood that have intensity values smaller than the smallest intensity value in the smaller neighborhood.

The output image is binary and is obtained by thresholding the output of the monotony operator. Selection of the specific threshold value, T , as well as, parameters m, n, k , and l is discussed in the next section.

The operator is applied in two steps, one that extracts predominantly horizontal structures and the other that extracts predominantly vertical structures. The elongated structures in arbitrary directions are obtained by combining the two results

by the logical OR operation, and their intersections are obtained by combining the two results by the logical AND operation. The centroids of obtained binary objects (typically smaller than 10 pixels) constitute the set of potential control points, Figure 2.

3.3 Parameter selection

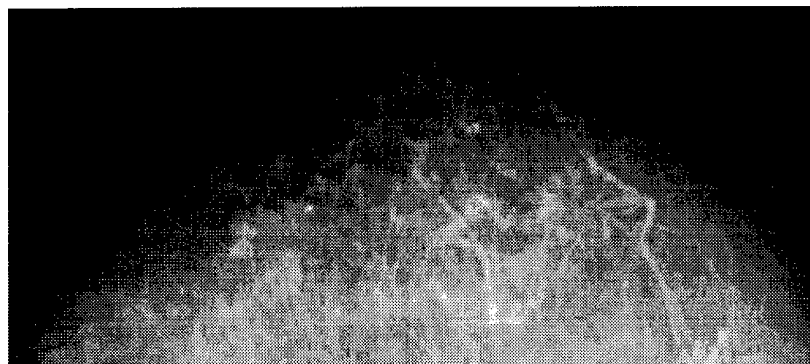
The parameters m, n, k, l employed by the monotony operator are selected such that in one case the operator detects predominantly horizontal structures and in the other predominantly vertical structures. The specifications are as follows:

- Detection of elongated vertical structures: $m = 11$, $n = 1$, $k = 3$, $l = 1$, and $T = 5$.
- Detection of elongated horizontal structures: $m = 1$, $n = 11$, $k = 1$, $l = 3$, and $T = 5$.

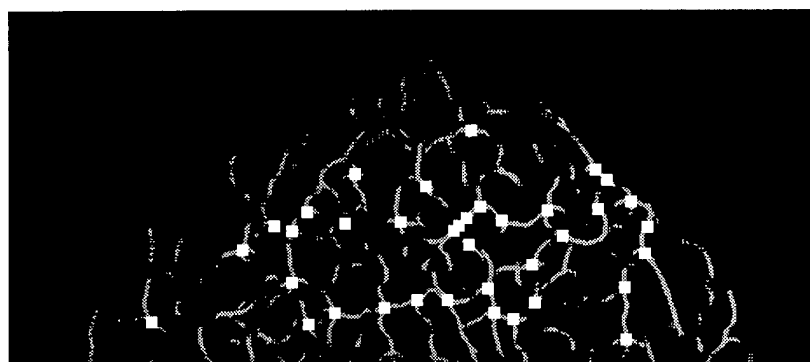
The parameters m (in the case of vertical structures) and n (in the case of horizontal structures) need to ensure incorporation of the complete width of an elongated structure. This, in turn, is determined by the resolution of input images and level of Gaussian blurring. Parameters k and l ensure incorporation of the central part of the profiles, which may be flat and few pixels wide, in blurred images. The selection of threshold, T , is guided by the objective of extracting only the central pixels. Considering the selection of specific values for m , n , k , l , the largest value is assigned to the central pixel and it is $(m \times n) - (k \times l) = 8$. If the operator is moved one location away from the central pixel, it generates value 5 or 6, depending on the profile shape. Since the objective is to extract only the central pixel and its immediate neighbors, the threshold is set to $T = 5$.

4 Establishing control point correspondence

Given two sets of potential control points, the objective of the correspondence algorithm is to identify the common subset and within it individual point correspondences. Most of the point pattern matching algorithms assume that the form of mapping relating the two sets of points is known; e.g., it may involve translation, rotation, or affine transformation. Since the relation between the two images in the case of mammograms is unknown, it is necessary to use a model-free matching algorithm.



(a)



(b)

Figure 2: An example of extracted elongated anatomical structures and detected potential control points: (a) original image, (b) extracted elongated structures and potential control points (shown enlarged).

4.1 Correspondence algorithm

The locations of corresponding points differ in their local coordinate systems; however, their relative positions within images are the same¹. Matching is attempted between a point $p(x_p, y_p)$ in the older mammogram and a point $q(x_q, y_q)$ in the newer mammogram only if point q lies in a region likely to provide a match². This region is determined using a pair of reference points $o(x_o, y_o)$ and $n(x_n, y_n)$ in the old and new mammogram, respectively. The correspondence between o and n is established accurately as described in Section 4.2. Point q is a match of point p if it lies in the neighborhood $k \times l$ centered at x_c, y_c determined by the intersection³ of the line

$$y = \frac{y_p - y_o}{x_p - x_o}(x - x_n) + y_n \quad (3)$$

and the circle

$$(x - x_n)^2 + (y - y_n)^2 = (x_o - x_p)^2 + (y_p - y_o)^2. \quad (4)$$

Values k and l are determined by expected differences between the two mammograms and in this work they are set to 60.

The correspondence is established using

- *Accumulator matrix* This matrix tallies votes for possible matches. Given a set of N_o potential control points in the older mammogram, M_o , and a set of N_n potential control points in the newer mammogram, M_n , the accumulator matrix is an $N_o \times N_n$, where entry $e(p, q)$ corresponds to points labeled p and q in M_o and M_n , respectively. Entry $e(p, q)$ is updated each time a possible match between points p and q is established.
- *Similarity criterion* This is the basis of establishing a match between a pair of points. The criterion measures similarity between the neighborhoods of two candidate points. Two types of criteria are considered based on (i) geometry of elongated structures and (ii) texture characteristics. The methods of characterization and the criteria are discussed in Section 4.3.

The entries of the accumulator matrix are formed in three steps:

- *Step 1: Location test* For each point in the old mammogram any potential control point in the new mammogram that satisfies the location criterion, i.e., it lies in the neighborhood $k \times l$ centered at (x_c, y_c) determined by Equations (3) and (4), is considered.

¹We consider the cases where the tissue is not twisted in either of the acquisitions.

²In principle, the proposed algorithm does not require limiting point locations, but constraining the search to image subregions significantly speeds up the processing time.

³There exist two points of intersection and the one of interest lies in the area of breast tissue since the reference point is the tip of the nipple.

- *Step 2: Matrix updating* If a matching criterion between a point p in the old mammogram and a point q in the new mammogram is satisfied, the accumulator entry $e(p, q)$ is updated. In the case that no match is established, return to Step 1; otherwise, proceed to Step 3.
- *Step 3: Processing the remaining points* If points $p(x_p, y_p)$ and $q(x_q, y_q)$ are matched in Step 2, then for each remaining point in the old mammogram, $r(x_r, y_r)$, a neighborhood of size $\Delta_2 \times \Delta_2$ of point (x'_c, y'_c) in the new mammogram is examined for a possible match. The coordinates (x'_c, y'_c) are determined by the intersection of the line

$$y = \frac{y_r - y_p}{x_r - x_p}(x - x_q) + y_q \quad (5)$$

and the circle

$$(x - x_q)^2 + (y - y_q)^2 = (x_r - x_p)^2 + (y_r - y_p)^2. \quad (6)$$

Each time a matching criterion is satisfied, the appropriate accumulator entry is updated. After considering all possible matches return to Step 1.

At the end of the matching process, the entries which correspond to true matches tend to have high values while the ones that correspond to wrong matches are small. The wrong matches are eliminated by thresholding. A good threshold value provides high values both for sensitivity (or the fraction of the established matches computed relative to the all possible matches) and specificity (or accuracy of the established matches). The emphasis of this work is on high specificity and, therefore, in some cases the number of established matches is relatively small.

Two approaches to thresholding the accumulator matrix were considered. The first one is a modification of the optimum thresholding algorithm described in [18]. The second one, used to obtain results discussed in Sections 5 and 6, uses the fixed threshold value of 80% of the maximum entry. The rationale for this value is discussed in Section 5.3.2.

After thresholding, the accumulator matrix is scanned and the most likely matches are determined. Scanning is done by first examining rows and then columns of the accumulator matrix. If there is more than one non-zero entry in a particular row, then the point in the new mammogram which corresponds to the largest accumulator entry is selected as the best match. The rest of the entries in the same row are set to zero. Columns are scanned in a similar fashion. After scanning, the accumulator matrix contains non-zero entries only at positions which correspond to the final matches.

4.2 Identifying reference points

Term reference points pertains to a pair of points whose correspondence is established accurately in a mammogram pair. The correspondence may be established manually or automatically and may utilize points intrinsic to the breast or defined by external markers used during the screening process. In principle, detection of reference points is independent of detection of potential control points. The reference points are used by the correspondence algorithm to identify possible matches among the two sets of potential control points.

The natural selection, i.e., a point existing in practically every mammogram and independent of view point, is the tip of the nipple, and in the following we describe the procedure used to identify automatically this point in two mammograms. Under normal screening conditions, the tip of the nipple lies approximately on the extrema point of the breast outline. In this work, this point is identified by finding the extrema of the Least Mean Square Error (LMSE) approximation of the breast outline by a quadratic function. The outline is initially determined by the optimum thresholding procedure, [18], applied to the blurred mammograms. The procedure was tested on 29 mammogram pairs, and visually it was confirmed that the reference points were identified consistently. It is pointed out that the same procedure was applied to both the medio-lateral and the cranio-caudal views without adjustments.

4.3 Similarity measures and criteria

4.3.1 Signatures

The objective of this approach is to capture local arrangements of dominant elongated structures in the neighborhood of potential control points. The algorithm elegantly and efficiently resolves the inherent compromise between accurate modeling and handling local distortions and yields considerably higher number of control points than the alternative described in [27]. The idea behind the approach can be best explained when considering the retino-cortical projection, i.e., a geometrical transformation described as a conformal mapping of the polar (ρ, θ) plane onto the Cartesian $(\log(\rho), \theta)$ plane. A point $z(\rho, \theta) = \rho e^{j\theta}$ in the retinal plane maps onto the point $w(u, v) = u + jv$ in the cortical plane as⁴

$$w = \log(z) = \log(\rho) + j(\theta + 2\pi k), \quad k = 0, \pm 1, \pm 2, \dots \quad (7)$$

The transformation may be implemented efficiently in a form of smart sensor, [21],[22].

⁴When taking into account only the principal branch of the logarithmic function, the representation of a cortical projection point is $u = \log(\rho)$, $v = \theta$.

Signature vectors are obtained from the cortical images by integrating the images along the ρ axis in intervals $\Delta\theta$. It is necessary that the sampling rates along the θ axis provide for local deviations from linearity; through experimentation the sampling rate is selected to be $\Delta\theta = 20^\circ$. The high values in a signature correspond to the elongated structures forming the intersection. Typically, a signature has 2-4 maxima. The locations of these maxima are encoded by thresholding the signature entries. Assuming that the length of an elongated structure encoded by the signature is l and its average width is w , the threshold value is selected to be $\frac{1}{2}(l \times w)$. Coefficient $\frac{1}{2}$ is due to the fact that the signature encompasses a wedge which covers approximately one half of the rectangular area that may be encompassed by the elongated structure. Thus, with $w = 4$, and $l = 10$, the threshold is set to 20.

The most fundamental issue in using the signatures is verifying that (i) control points are characterized by unique signatures, and (ii) signatures are preserved in subsequent screenings. Since the signatures in effect capture the orientation of the elongated structures it is necessary to show that the two properties hold for the orientations. Considering that there is no "ground truth" corresponding to mammograms, we have studied the distribution of orientations of elongated structures at potential control points in mammogram pairs. Ideally, the distributions should be uniform in order to provide for signature uniqueness and the corresponding distributions should be identical to provide for signature preservation. The analysis of 31 mammogram pairs has shown that both conditions are satisfied.

Similarity criterion

Two signatures $S_1 = [s_k^1]$ and $S_2 = [s_k^2]$, $k = 1, 2, \dots, \frac{360^\circ}{\Delta\theta}$ match if $r \geq t$, where

$$r = S_1 S_2 = \sum_{k=1}^{360^\circ/\Delta\theta} s_k^1 s_k^2, \quad (8)$$

and t is the threshold. Since the signatures are binary vectors, the similarity criterion given by Equation (8) counts the number of 1's that appear at the same positions in two signature vectors. Assuming that at least two parts of elongated structures are captured by a signature, we select $t = 2$.

4.3.2 Texture characterization

Texture may be modeled in various ways, e.g., using descriptors derived from co-occurrence matrices, fractal measures, or Markov random field descriptors. Based on the work of other researchers in digital mammography, e.g., [8], [17], [20], we have adopted Laws measures as the descriptors of textural properties.

Laws' texture measures are formed in two steps [12]. In the first step, correlation between small masks (usually of size 5×5) and the image pixels of interest is performed. Each of the masks is designed to respond to a specific texture characteristic

(e.g., gray level, ripple, edge, spot, or their combinations). The result of the correlation between the input image and a particular mask is referred to as the feature plane of that mask. In the second step, macrostatistics are computed over a larger neighborhood (15×15) in the feature planes. These macrostatistics typically involve standard deviation or its approximation (such as the absolute average). The macrostatistics are referred to as the Laws' texture features or texture energy features. We have used the entire set of 15 invariant⁵ features to form feature vectors.

Similarity criterion

If a point p in the old mammogram and a point q in the new mammogram are characterized by the Laws feature vectors LV_p and LV_q , respectively, then q is a match for p if

$$\frac{\|LV_p - LV_q\|}{\|LV_p\|} = \frac{\sqrt{\sum_{i=1}^{15} (f_i^p - f_i^q)^2}}{\sqrt{\sum_{i=1}^{15} (f_i^p)^2}} < t, \quad (9)$$

where $\|\cdot\|$ denotes the length of the vector and t is the threshold. Thus, the match between vectors LV_p and LV_q is established if the Euclidean distance between vectors LV_p and LV_q is less than some prespecified fraction t of the length of LV_p . We point out that the criterion is robust with respect to t and stable results are obtained for t in the range $[0.25 : 0.4]$; we use $t = .3$.

5 Evaluation: Synthetic images

Evaluation of the proposed algorithms was initially carried out using synthetic images designed to evaluate independently the algorithm for extraction of elongated structures and the correspondence algorithm. In the following we first summarize characteristics of images used in the study and then describe the evaluation protocols.

5.1 Material selection

The synthetic images were designed to evaluate the algorithms' sensitivity to variations in orientations of elongated structures, their contrast relative to background, background characteristics, noise and parameters employed by the algorithms. The underlying model in generating the elongated structures was the Poissonian model, which randomly distributes linear structures [16]. In practical terms, the Poissonian model assumes a network of intersecting straight lines with random positions and orientations. The linear structures were deformed using the deflected elastic beam model, [24].

⁵The term invariant refers to the gray level; for details see [12].

5.2 Evaluation: detection algorithm

5.2.1 Experiment description

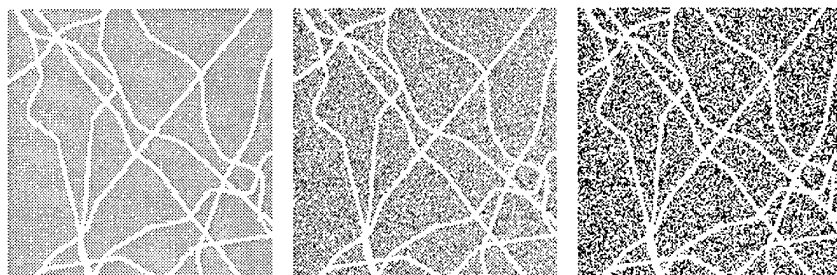
The objective of these experiments was to determine the effects of width of linear structures, their contrast relative to background, background type, and noise on algorithm's performance. Ten patterns of elongated structures were generated, and each was used to form eight images by changing the width and the intensity of structures. Four of the images contained structures of width 1, 3, 5, and 7 pixels and intensity 255. Four images had structures of constant width of 3 pixels and intensity 100, 150, 200 and 220, respectively. The structures were superimposed on the following backgrounds

- *Case 1* Two types of backgrounds were used: (1) five images with noisy background formed by the Gaussian random process of mean $\mu = 200$ and standard deviations 10, 20, 50, 80, 200, respectively, and (2) ten images with texture backgrounds selected from [1]. In both cases, images were quantized to 256 gray levels. These backgrounds combined with eight patterns formed $8 \times 10 \times 15$ images. Examples of synthetic images generated in this manner are shown in Figures 3 (a) and 3(b).
- *Case 2* In this case, backgrounds were obtained using patches from mammograms. A total of 25 samples was used: 10 radio-lucent, 10 mixed, and 5 radio-opaque (for characteristics of each, see Section 6.1). These backgrounds combined with the eight patterns formed $8 \times 10 \times 25$ images. Examples of obtained images are shown in Figure 3(c).

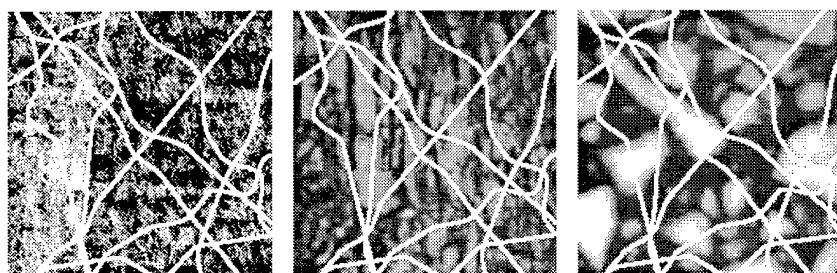
5.2.2 Results

Case 1

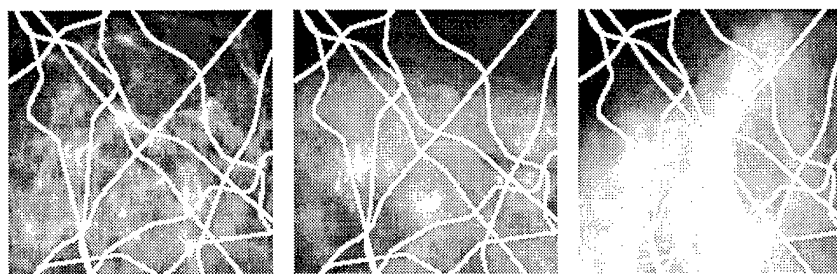
In this case, the receiver operating characteristics (ROC) curves were used. A ROC curve is determined by a set of points (FPF, TPF) which are obtained using different thresholds [14], [15]; TPF and FPF stand for true positive fraction and false positive fraction, respectively. In this experiment, the TPF is the fraction of pixels belonging to the linear structures that are correctly identified by the algorithm as "positives". The FPF is computed as $FPF = 1 - TNF$, where TNF (true negative fraction) is the fraction of pixels that do not belong to linear structures that are correctly identified by the algorithm as "negatives". The TPF and TNF represent sensitivity and specificity, respectively. For the modified monotony operator, the range of possible thresholds is [1:8], which means that each ROC curve was determined by computing eight points, i.e., eight TPF and FPF values.



(a)



(b)



(c)

Figure 3: Examples of images used in evaluation of the detection algorithm: background and superimposed elongated patterns. (a) Case 1-noisy backgrounds (b) case 1-texture backgrounds, and (c) case 2-mammogram backgrounds.

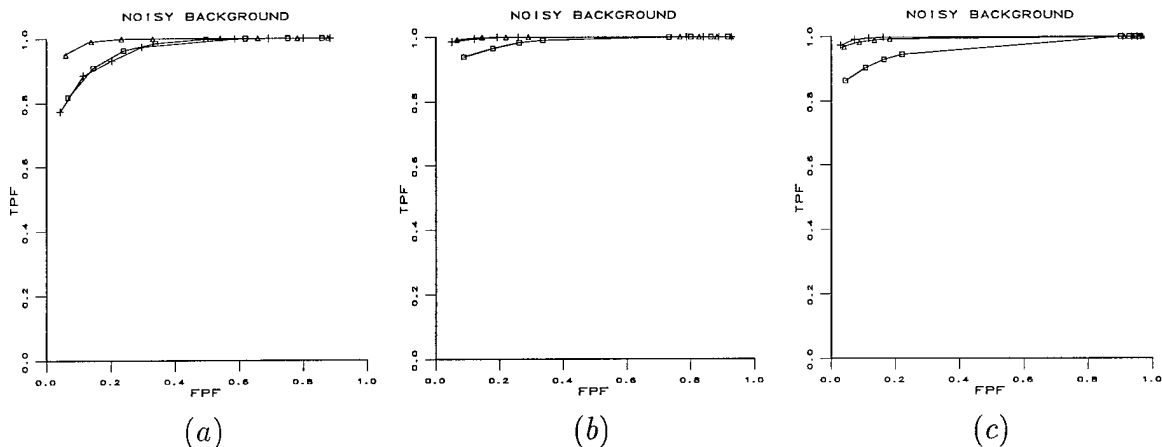


Figure 4: ROC curves generated when considering noisy backgrounds and Gaussian smoothing is performed in (a) one, (b) two, and (c) three passes. Elongated structures were 1 (square), 3 (triangle), and 5 (cross) pixels wide.

In the following we discuss the algorithm's sensitivity to the width of elongated structures, level of Gaussian blurring and type of background. The details of other experiments are described in [26], including the experiments on sensitivity to contrast which have shown that the operator is capable of detecting very low contrast roof edges and its performance is satisfactory as long as the elongated structures have higher intensity than the local average of the background (thus, for noisy backgrounds, intensity 220 of elongated structures yielded good results). In the current work, we show two sets of ROC curves. The first set, Figure 4, shows the algorithm's performance on noisy background, and the second set, Figure 5, shows the algorithm's performance on texture backgrounds. In both cases, the effects of Gaussian blurring on the results are shown, indicating that the three-pass Gaussian filtering used on archived film yields good results. Different ROC curves were obtained by changing the width of the elongated structures.

From these results it follows that the operator using parameters discussed in Section 3.3 provides consistent results for a range of structure widths. Poorer performance of the algorithm for one-pixel wide structures is attributed to Gaussian filtering, which tends to eliminate the thin structures. Figures 4 and 5 also show that : (i) the sensitivity and specificity are not affected by the type of background (noisy or textured) and (ii) the best compromise between sensitivity and specificity of the algorithm is obtained for the threshold value of five, which is in agreement with theoretical

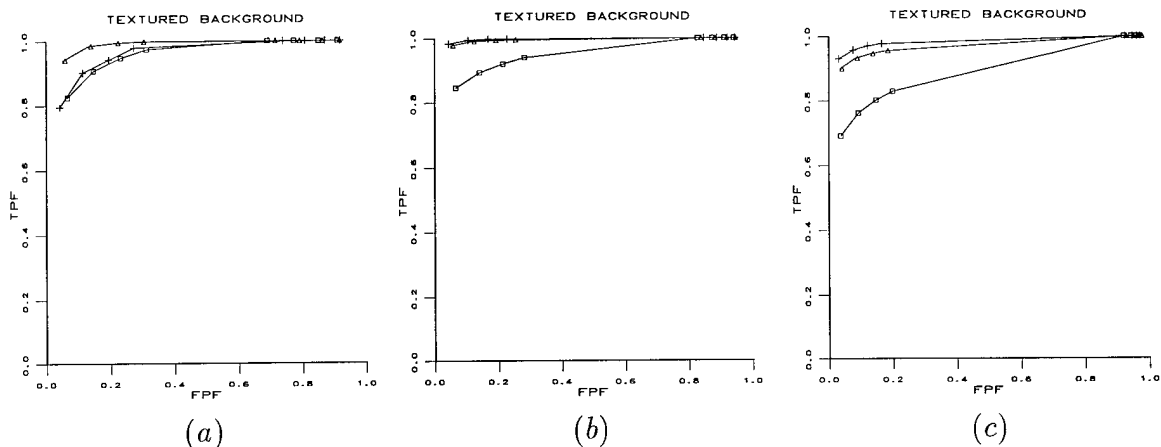


Figure 5: ROC curves generated when considering textured backgrounds and Gaussian smoothing is performed in (a) one, (b) two, and (c) three passes. Elongated structures were 1 (square), 3 (triangle), and 5 (cross) pixels wide.

considerations discussed in Section 3.3.

Case 2

In this case, the background images contained linear structures generic to mammograms; therefore, the FPF could not be determined, which precluded using the ROC curves. However, the TPF can be determined and the algorithm was evaluated only with respect to sensitivity. The results are summarized in Table 1 for threshold $T = 5$.

The following conclusions can be drawn from these experiments. The algorithm is insensitive to mammogram type. The algorithm is also insensitive to change in the structure's width for structures wider than one pixel. By comparing the results from the two experiments, it is evident that the algorithm's performance is not affected by the type of the background.

5.3 Evaluation: correspondence algorithm

Due to the fact that there is no objective way of simulating appropriate texture patterns, this evaluation includes only the signature-based approach. The purpose of the tests is to determine the robustness of the correspondence algorithm with respect to parameters employed by the algorithm.

| lucent | mixed | dense | width | level of smoothing |
|--------|-------|-------|-------|--------------------|
| 99.5 | 99.8 | 99.9 | 1 | 1 |
| 98.0 | 98.9 | 99.5 | 1 | 2 |
| 85.0 | 89.0 | 83.4 | 1 | 3 |
| 99.9 | 99.9 | 99.9 | 3 | 1 |
| 99.9 | 99.9 | 100 | 3 | 2 |
| 98.0 | 98.0 | 98.0 | 3 | 3 |
| 97.8 | 98.2 | 97.9 | 5 | 1 |
| 99.9 | 100 | 100 | 5 | 2 |
| 99.2 | 99.2 | 99.5 | 5 | 3 |

Table 1: True positive fraction (in percent) computed for Case 2 using $T = 5$.

5.3.1 Experiment description

In this experiment, we have tested the performance of the matching algorithm with respect to Δ_1 and Δ_2 . The results are summarized in the form of ROC curves. For each test, a set of 100 image pairs was generated. Each image pair comprised an original and a corresponding deformed pattern. Then, the matching algorithm was run and, depending on the threshold, different points on the ROC curve were generated. Each ROC curve was determined by ten points, which were obtained by thresholding the accumulator array at 0, 10, ..., 90 percent of the maximum value in the array.

5.3.2 Results

Two sets of ROC curves, Figure 6, summarize the results. The following conclusions can be drawn from the curves. Figure 6(a) indicates that the algorithm is insensitive to change of parameter Δ_1 . Also, the performance is not affected by Δ_2 (Figure 6(b)). Note that combining, e.g., Δ_1 of 100 and Δ_2 of 100 produces a search area of 200×200 pixels which covers practically the whole image. This implies that each potential control point in one image is matched against the whole other image and the algorithm's performance remains stable. It should be noted that both good specificity and sensitivity of the algorithm can be achieved, since for TPF of about 97% FPF is smaller than 5%. Figure 6 shows that the accumulator matrix threshold can always be determined to accommodate for high specificity, and thresholds of 80% of the maximum accumulator entry are always a good selection. We have also investigated the algorithm's performance in terms of maximum distortion which is a parameter in

the deflected elastic beam model. In general, the algorithm's performance decreases with the increase in distortion, but even for high levels of distortion a working point with low FPF can be achieved for $T=5$, [26].

6 Evaluation: Archived film

6.1 Material selection

Each case used in the study was represented by mammograms of two screenings in standard cranio-caudal and medio-lateral views. In order to determine the effects of breast tissue characteristics on the method's performance, the selected cases reflected three groups

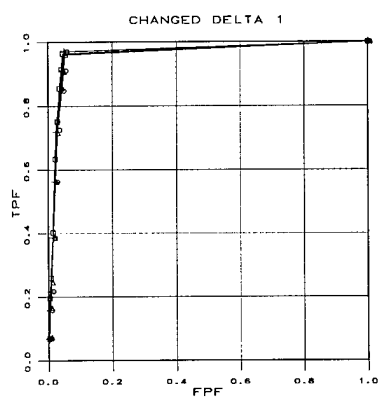
- *Radio-opaque/dense cases* (type labeled DY by Wolfe, [32]). These images are very bright and low-contrast, have no distinctive regions, and there is very little structure to their texture patterns. The human visual system is limited the most in evaluating these images due to lack of ability to differentiate details at these intensity levels.
- *Radio-lucent cases* (N1 type according to Wolfe, [32]). These patterns have structure and contrast.
- *Mixed cases* (labeled P1 and P2 by Wolfe, [32]). These cases are characterized by two types of texture patterns, the very bright dense patterns (characteristic of DY types) and structural texture patterns (characteristic of N1 type).

The algorithms developed in this study were tested on 31 mammogram pairs. Based on tissue characteristics, 16 pairs had predominantly radio-lucent characteristics, 10 pairs contained mixed tissue types, and 5 pairs had predominantly radio-opaque characteristics. Spatial resolution varied from $.1mm$ to $.5mm$, and only mammogram pairs showing visual similarity were used in the study.

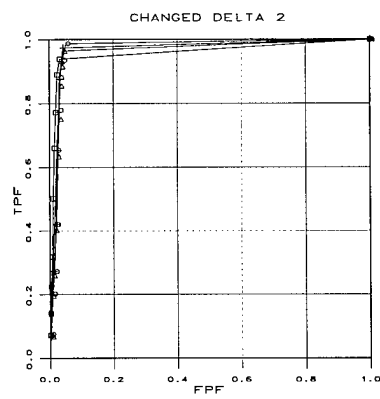
6.2 Experiment description

Due to lack of ground truth regarding control points, the evaluation was performed by visual inspection. Validation was made independently by three unbiased observers, i.e., observers who had no knowledge about the algorithm: an experienced radiologist and two untrained observers.

The validation experiment was conducted as follows. Each observer was shown a mammogram pair, where the older mammogram contained highlighted points selected by the algorithm. The observer was then asked to identify the corresponding



(a)



(b)

Figure 6: (a) ROC curves generated for Δ_1 20 (square), 40 (triangle), 60 (cross), and 100 (circle); (b) ROC curves generated for Δ_2 20 (square), 40 (triangle), 60 (cross), and 100 (circle).

points in the newer mammogram. Observers were also asked to assign a "level of confidence" to the established matches. Namely, if a match was established with a high degree of certainty, the confidence level was "sure"; if a match was determined with a lower degree of certainty, the confidence level was "not sure"; in the case when no match could be assigned, a point was labeled "don't know". No time limit was imposed upon observers during the validation experiment. The observers were trained and encouraged to use enhancement techniques in the experiment (e.g., histogram equalization).

6.2.1 Analysis of consistency of human performance

The basic issue in involving human subjects in an experiment is establishing consistency of their performance and presence of bias. In order to validate the results used in the study, two untrained observers were presented three pairs of mammograms eight times. The older mammogram in each pair contained highlighted control points determined by the computer. In each presentation the observers recorded corresponding points in the new mammogram. In the first four presentations, the observers were shown the identical mammogram pair. In the fifth, sixth, and seventh presentations, the mammograms were flipped around x axis, around y axis, and around both x and y axes, respectively. In the eighth presentation only a portion of the mammograms that contained points of interest were shown, i.e., the images contained no visible breast outline. This was done to check if the human perception of control points is affected by landmarks such as image borders and/or breast outline. The obtained results were then compared using statistical tests.

The three mammogram pairs were classified as "easy", "not-so-easy", and "difficult", corresponding to radio-lucent, mixed, and radio-opaque, respectively. Statistical analysis was used to determine if the two tested sets of data (or two populations) had the same mean. The statistical hypotheses were as follows:

H_0 : *The two populations have equal means*

H_1 : *The two populations have different means.*

The objective was to determine if there is enough statistical evidence to reject the null hypothesis.

We have utilized five statistical tests, one multivariate and four univariate. Based on the fact that data is two-dimensional,⁶ multivariate tests are more suitable. The univariate tests were adapted to test multidimensional data. The adaptation was done as follows. Each of the random variables x and y (the coordinates of the points determined by the observers) was tested separately utilizing univariate tests. The null hypothesis was rejected if either of the hypotheses for univariate data was rejected.

⁶Each point contains two coordinates, x and y , which are considered to be independent random variables.

| | observer 1 | observer 2 |
|------------|--------------|--------------|
| observer 1 | (100,100,66) | (66,20,100) |
| observer 2 | (66,20,100) | (100,83,100) |

Table 2: Result of T^2 test.

| | observer 1 | observer 2 |
|------------|--------------|---------------|
| observer 1 | (100,83,100) | (83,66,100) |
| observer 2 | (83,66,100) | (100,100,100) |

Table 3: Result of t-test (equal variances).

Consequently, if the univariate tests are performed with the levels of significance α_x and α_y , for x and y respectively, the overall level of significance α for the two-dimensional data is

$$\alpha = 1 - (1 - \alpha_x)(1 - \alpha_y). \quad (10)$$

In the study we have utilized the following tests: the Hotelling T^2 test, two t-tests, the Kolmogorov-Smirnov test, and the rank sum test. All of these are standard and can be found in books of mathematical statistics (e.g., [7], [13]). The results of the five statistical tests are summarized in Tables 2-6. Each entry in the table is represented by a triplet of numbers (X, Y, Z) which represent the percentage of points in agreement between tested populations for “easy”, “not-so-easy”, and “difficult” mammogram pairs, respectively. Thus, the entry (*observer 1, observer 2*)=(66,20,100) in Table 2 means that observer 1 agreed with observer 2 in 66% of points when considering “easy” mammogram pair, in 20% of points when considering “not-so-easy” mammogram pair, and in 100% of points when considering “difficult” mammogram pair using the Hotelling T^2 test. Similarly, the entry (*observer1, observer1*) = (100,83,100) in Table 3 means that observer 1 was always consistent when considering “easy” and “difficult” mammogram pairs and was consistent in 83% of points when considering “not-so-easy” mammogram pair using the univariate t-test (variances equal and unknown). The number of points considered in the “easy”, “not-so-easy”, and “difficult” mammogram pairs was 6, 6, and 3, respectively.

Several conclusions can be drawn from Tables 2 - 6. The Wilcoxon-Mann-Whitney rank sum test indicates that an observer has always consistent perception of control points regardless of the mammogram type. However, the t -tests indicate that this is not always the case. Since the Wilcoxon-Mann-Whitney rank sum test does not

| | observer 1 | observer 2 |
|------------|--------------|--------------|
| observer 1 | (100,100,66) | (66,16,100) |
| observer 2 | (66,16,100) | (100,83,100) |

Table 4: Result of t-test (unequal variances).

| | observer 1 | observer 2 |
|------------|---------------|---------------|
| observer 1 | (100,100,100) | (100,100,100) |
| observer 2 | (100,100,100) | (100,100,100) |

Table 5: Result of the Kolmogorov-Smirnov test.

| | observer 1 | observer 2 |
|------------|---------------|---------------|
| observer 1 | (100,100,100) | (100,50,100) |
| observer 2 | (100,50,100) | (100,100,100) |

Table 6: Result of the Wilcoxon-Mann-Whitney rank sum test.

| | algorithm | radiologist | observer 1 | observer 2 |
|-------------|-----------|-------------|------------|------------|
| algorithm | - | 91 | 71 | 75 |
| radiologist | - | - | 83 | 82 |
| observer 1 | - | - | - | 78 |

Table 7: Result of validation when using signatures and considering only “sure” points. Entries represent the percentage of points “in agreement”.

impose any restrictions upon data and the t -based tests assume normal distributions, we conclude that the data generated by the observers are not probably normally distributed. The T^2 test also makes the assumption about the normal distribution of data which means that this test may not be also appropriate. Kolmogorov-Smirnov test has never failed, which means that in the statistical sense there is not enough evidence that the tested populations ever came from different distributions. The Wilcoxon-Mann-Whitney rank sum test failed in 50% of the cases when two populations generated by two observers were tested for the “not-so-easy” mammograms.

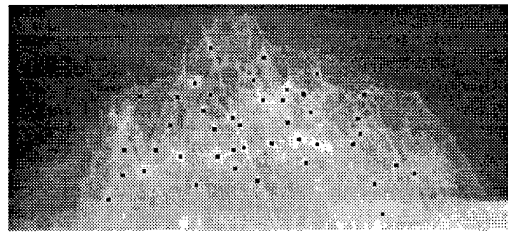
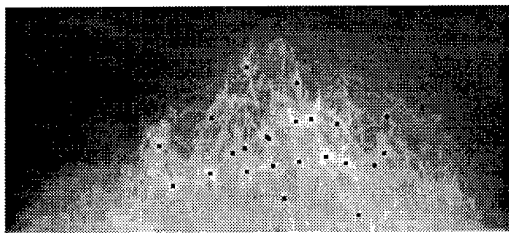
In summary, regardless of the mammogram type the human perception is consistent. This consistency is our justification of using observers in evaluation of the algorithms’ performance on archived film. In addition, the human perception is not affected by landmarks such as the breast outline.

6.3 Results: signatures

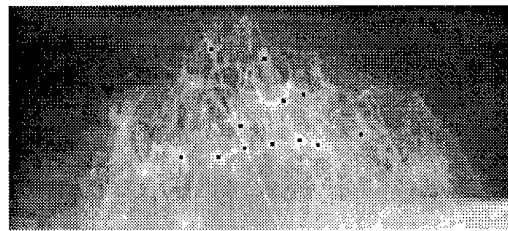
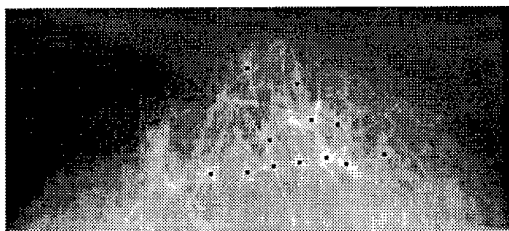
Based on the spatial resolution, the algorithm and an observer were considered to be in agreement if the distance between the two points was less than 10 pixels. The algorithm established correspondence between 11 pairs of points on average; this is in contrast to the earlier work, [27], that was establishing correspondence between 5 pairs on average. Figures 7(b) and 8(b) show two examples of detected control points using signatures.

Table 7 summarizes the validation results when only points labeled “sure” were considered. Each entry in the table represents the percentage of points “in agreement”. For example, the entry $(algorithm, radiologist) = 91$ indicates that 91% of “sure” control points determined by the radiologist are “in agreement” with corresponding points determined by the computer. Table 8 includes both “sure” and “not sure” points.

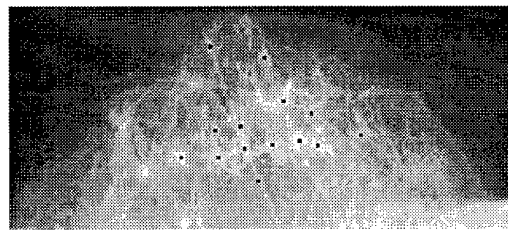
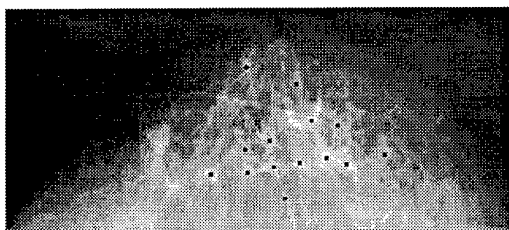
In order to test the sensitivity of the algorithm to breast tissue characteristics, results pertaining to 16 pairs of mammograms exhibiting a high degree of radiolucency were analyzed separately. The validation results in this case indicate that the



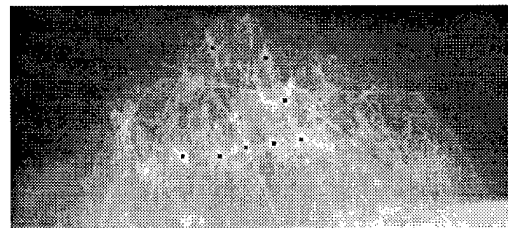
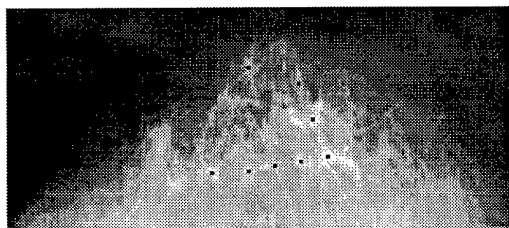
(a)



(b)

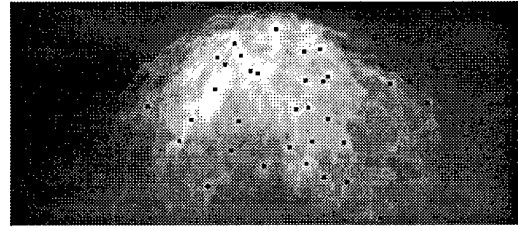
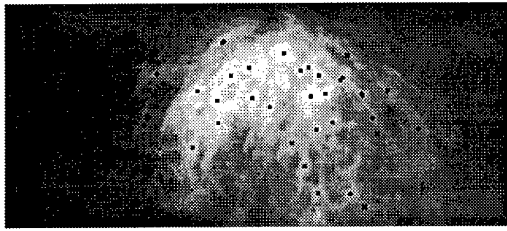


(c)

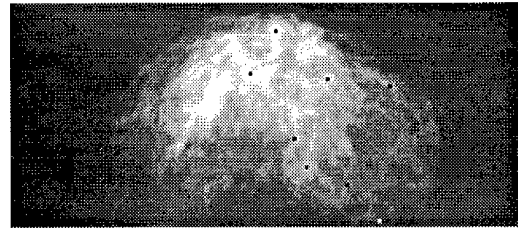
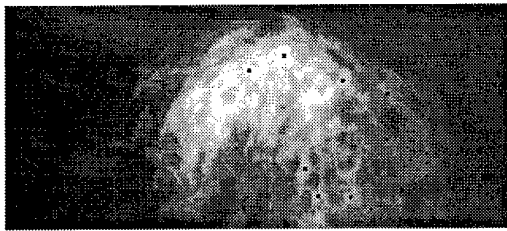


(d)

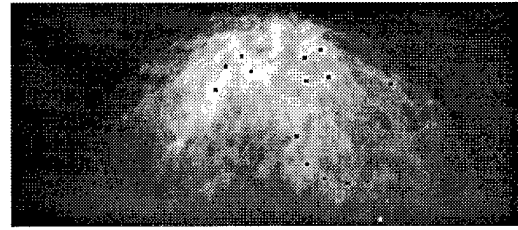
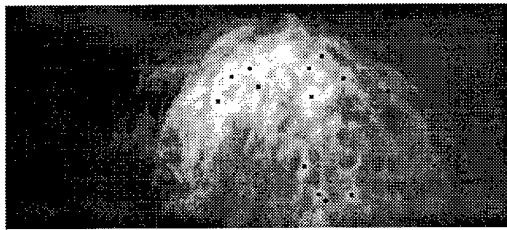
Figure 7: Extracted points on a pair of radio-lucent mammograms (a) potential control points, (b) matched points using signatures, (c) matched points using Laws' descriptors, (d) result of combining signatures and texture descriptors.



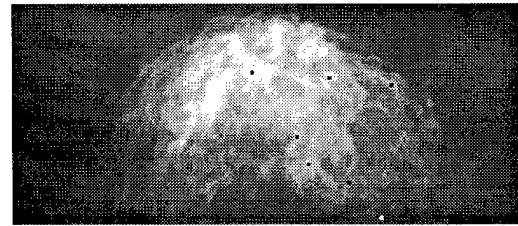
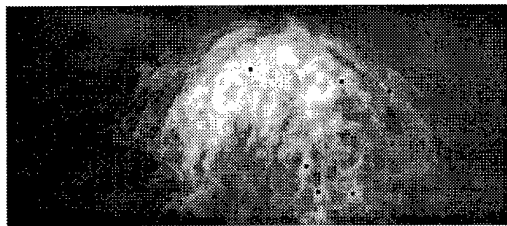
(a)



(b)



(c)



(d)

Figure 8: Extracted points on a pair of mixed tissue type mammograms (a) potential control points, (b) matched points using signatures, (c) matched points using Laws' descriptors, (d) result of combining signatures and texture descriptors.

| | algorithm | radiologist | observer 1 | observer 2 |
|-------------|-----------|-------------|------------|------------|
| algorithm | - | 86 | 72 | 69 |
| radiologist | - | - | 78 | 77 |
| observer 1 | - | - | - | 67 |

Table 8: Result of validation when using signatures and considering both “sure” and “not sure” points. Entries represent the percentage of points “in agreement”.

| | algorithm | observer 1 | observer 2 |
|------------|-----------|------------|------------|
| algorithm | - | 93 | 81 |
| observer 1 | - | - | 90 |

Table 9: Result of validation of the algorithm based on Laws texture measures.

conclusions drawn for the whole set of mammograms remain valid for this subset as well, thus showing that the algorithm’s performance does not depend on mammogram characteristics.

6.4 Results: Laws’ texture measures

The matching algorithm utilizing Laws’ measures was run on the set of 31 mammogram pairs, and the established matches were validated by visual inspection involving two observers. The algorithm establishes a relatively large number of matches (14 points on average) with good accuracy (over 81% of the established pairs were evaluated to be correct). Two examples of established matches are shown in Figures 7(c) and 8(c). The validation results are summarized in Table 9.

In the search for more accurate matching we have combined signatures and Laws texture features by simultaneously comparing both signatures and Laws feature vectors and tallying the votes for only those pairs that satisfy both the signature (Equation (8)) and Laws vector similarity criterion (Equation (9)). Two examples of established matches are shown in Figures 7(d) and 8(d). We ran the algorithm on 31 mammogram pairs and the results are as follows. It established on average 10 pairs of points per mammogram pair with the accuracy of at least 90%). The results are summarized in Table 10.

| | algorithm | observer 1 | observer 2 |
|------------|-----------|------------|------------|
| algorithm | - | 95 | 90 |
| observer 1 | - | - | 95 |

Table 10: Result of validation of the algorithm based on Laws texture measures and signatures.

7 Conclusions

The focus of this work is on identifying landmarks and establishing their correspondence in temporal pairs of mammograms. The motivation of the work is automation of analysis of mammogram sequences. Two algorithms were developed and tested. The first algorithm identifies potential control points in mammograms, and it was evaluated on synthetic images and was shown to be robust to changes in width of elongated structures, background complexity, and presence of noise. The second algorithm establishes correspondence between potential control points, and it was evaluated both on synthetic images and archived film. One of the major challenges of the work was developing evaluation protocols because no “ground truth images” exist. The bulk of the evaluation involved human subjects, and it was necessary to evaluate reliability of performance of an observer and to quantify consistence of the observer’s visual perception of control points selected by the algorithm. The study has shown that visual perception of the control points selected by the algorithm is consistent for a single observer. The most important conclusions are that there is the agreement in 90% of cases between the algorithm and the medical expert and that the algorithm’s performance is not affected by the mammogram type. Both of the criteria developed for measuring similarity of potential control points yield good results, and the most reliable results are obtained by combining signature and texture-based approaches.

Our present research focuses on using the control points to define corresponding regions. The appropriate region size and geometry are under study. The ultimate test of performance will be in the next stage of development when the complete system performance, including characterization of corresponding regions, is compared to results of biopsies.

References

- [1] P. Brodatz, *Textures: a Photographic Album for Artists and Designers*, Dover Publications, 1966.
- [2] D. Brzakovic and M. Neskovic, "Mammogram Screening Using Multiresolution-based Image Segmentation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 6, pp. 1437-1459, 1993.
- [3] P.J. Burt, "Fast Filter Transforms for Image Processing," *Computer Graphics and Image Processing*, vol. 16, 1981.
- [4] N. Cerneaz and M. Brady, "Finding Curvi-Linear Structures in Mammograms," Proc. of the First International Conference, CVRMed '95, Nice, France, pp. 372-382, 1995.
- [5] R.N. Dixon and C.J. Taylor, "Automated Asbestos Fiber Counting," *Inst. Phys. Conf. Series*, vol. 44, pp. 178-185, 1979.
- [6] W. Hand, J.L. Semmlow, L.V. Ackerman, and F.S. Alcorn, "Computer Screening of Xeromammograms: A Technique for Defining Suspicious Areas of the Breast," *Computers and Biomedical Research*, vol. 12, pp. 445-460, 1979.
- [7] G.K. Kanji, "100 statistical Tests," SAGE Publications, 1993.
- [8] W.P. Kegelmeyer, Jr., "Evaluation of Stellate Lesion Detection in a Standard Mammogram Data Set," in *State of the Art in Digital Mammographic Image Analysis*, Eds. K.W. Bowyer and S. Astley, World Scientific, Singapore, pp. 262-279, 1994.
- [9] H. Kobatake, "Detection of Spicules on Mammogram Based on Skeleton Analysis," *IEEE Transactions on Medical Imaging*, vol. 15, no. 3, pp. 235-245, June 1996.
- [10] R. Kories and G. Zimmerman, "A Versatile Method for the Estimation of Displacement Vector Fields from Image Sequences," Proc. of Workshop on Motion: Representation and Analysis, 1986.
- [11] T. Lau and W.F. Bischof, "Automated Detection of Breast Tumors Using the Asymmetry Approach," *Computers and Biomedical Research*, vol. 24, pp. 273-295, 1991.

- [12] K.I. Laws, Textured Image Segmentation, Ph.D Dissertation, University of Southern California, January 1980.
- [13] E.L. Lehmann, *Testing Statistical Hypotheses*, Wiley, New York, 1986.
- [14] C.E. Metz, "ROC Methodology in Radiographic Imaging," *Investigative Radiology*, vol. 21, pp. 720-733, 1986.
- [15] C.E. Metz, "Statistical Analysis of ROC Data in Evaluating Diagnostic Performance," in *Multiple Regression Analysis: Applications in Health Sciences*, Medical Physics Monograph, Eds. D.E. Helbert and R.H. Myers, vol. 13, pp. 365-384, 1986.
- [16] R.E. Miles, "A Various Aggregates of Random Polygons Determined by Random Lines in a Plane," *Adv. in Math*, vol. 10, pp. 256-290, 1973.
- [17] P. Miller and S. Astley, "Automated Detection of Breast Asymmetry Using Anatomical Features," in *State of the Art in Digital Mammographic Image Analysis*, Eds. K.W. Bowyer and S. Astley, World Scientific, Singapore, pp. 247-261, 1994.
- [18] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-9, no. 1, pp. 62-66, 1979.
- [19] T. Parr, S. Astley, and C. Boggis, "The Detection of Stellate Lesions in Digital Mammograms," Proc. of the 2nd International Workshop on Digital Mammography, pp. 231-239, York, England, July 1994.
- [20] M. Sallam and K. Bowyer, "Registering Time Sequences of Mammograms Using a Two-Dimensional Image Unwarping Technique," Proc. of the 2nd International Workshop on Digital Mammography, pp. 121-130, York, England, July 1994.
- [21] G. Sandini and P. Dario, "Active Vision Based on Space-Variant Sensing," *Proc. of the Fifth International Symposium on Robotics Research*, Tokyo, Japan, 1989.
- [22] G. Sandini and V. Tagliasco, "An Antropomorphic Retina-Like Structure for Scene Analysis," *Computer Graphics and Image Processing*, vol. 14, pp. 362-372, 1980.

- [23] J.L. Semmlow, A. Shadagopan, L.V. Ackerman, W. Hand, and F.S. Alcorn, "A Fully Automatic System for Screening Mammograms," *Computers and Biomedical Research*, vol. 13, pp. 350-362, 1980.
- [24] I.S. Sokolnikoff, *Mathematical Theory of Elasticity*, McGraw-Hill, New York, 1956.
- [25] E.A. Stamatakis, A.Y. Cairns, I.W. Ricketts, C. Walker, P.E. Preece, and A.J. Thompson, "A Novel Approach to Aligning Mammograms," Proc. of the 2nd International Workshop on Digital Mammography, pp. 355-364, York, England, July, 1994.
- [26] N. Vujovic, Registration of Time-Sequences of Random Textures with Applications to Mammogram Followup, Ph.D. Dissertation, Lehigh University, 1997.
- [27] N. Vujovic and D. Brzakovic, "Establishing Correspondence Between Pairs of Mammograms," in press, *IEEE Trans. on Image Processing*, 1997.
- [28] N. Vujovic and D. Brzakovic: "An Approach to Putting Mammogram Sequences into Correspondence," submitted to *IEEE Trans. on Medical Imaging*.
- [29] F. Yin, M.L. Giger, K. Doi, C.E. Metz, C.J. Vyborny, and R.A. Schmidt, "Computerized Detection of Masses in Digital Mammograms," *Medical Physics*, vol. 18, pp. 955-963, 1991.
- [30] F.F. Yin, M.L. Giger, K. Doi, C.J. Vyborny, and R.A. Schmidt, "Computerized Detection of Masses in Digital Mammograms: Automated Alignment of Breast Images and Its Effect on Bilateral-Subtraction Technique," *Medical Physics*, vol. 21, no. 3, pp. 445-452, March 1994.
- [31] F. Winsberg, M. Elkin, J. Macy, V. Bordaz, and W. Weymouth, "Detection of Abnormalities in Mammograms by Means of Optical Scanning and Computer Analysis," *Radiology*, vol. 89, pp. 211-215, 1967.
- [32] J.N. Wolfe, "Breast Patterns as an Index of Risk for Developing Breast Cancer," *American Journal of Roentgenology*, vol. 126, pp. 1130-1139, 1976.